DOCUMENT RESUME

ED 169 121 TM 008 608

AUTHOR Perry, Barbara Lee
TITLE The Effects of Three Methods of Interactive Programmed Training on the Accuracy of Classroom Assessors.
PUB DATE [Jan 78]
NOTE 22p.; Paper presented at the Annual Meeting of the Georgia Educational Research Association (2nd, Atlanta, Georgia, January 13-14, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classroom Observation Techniques; Elementary School Teachers; Elementary Secondary Education; Inservice Programs; *Inservice Teacher Education; Observation; *Participant Satisfaction; *Peer Evaluation; Programed Instruction; Programed Materials; Secondary School Teachers; *Teacher Evaluation; *Teaching Methods; Video Cassette Systems
IDENTIFIERS Classroom Procedures Assessment Instrument; Interpersonal Skills Assessment Instrument

ABSTRACT

Elementary and secondary school teachers were trained in the use of two observation instruments for teacher evaluation--the Classroom Procedures Assessment Instrument and the Interpersonal Skills Assessment Instrument. Training was conducted in workshops which used a series of ten videocassette sequences showing teachers working with students. Accompanying printed materials presented questions and answers which explained the meaning of specific indicators on the two instruments; these questions helped the trainee rate the videocassette teachers' behavior. Groups viewed the videocassette and then participated in either: (1) oral discussion with feedback from an instructor; (2) directed discussion with an instructor; (3) self-tutorial using a manual, without discussion; or (4) no treatment. The raters were also categorized according to the grade level they taught (K-5 or 6-12). Their attitude toward the training they received and their performance in rating teacher behavior in a 30-45 minute videocassette were studied. Analysis of variance indicated that type of training did affect performance in rating; use of written responses with either oral discussion or printed feedback was more successful than oral discussion only. Practice and rating in one's own grade level diminished these effects. No significant attitudinal differences were found. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED169121

THE EFFECTS OF THREE METHODS OF INTERACTIVE PROGRAMMED TRAINING ON THE ACCURACY OF CLASSROOM ASSESSORS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Barbara Perry

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Barbara Perry

College of Education

University of Georgia

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The accountability movement in education during the 1960s has encouraged attempts to study teacher effectiveness or teacher competency. Local communities have pressured school systems to justify their need for more tax dollars. State legislatures have acted on this movement by enacting accountability legislation, such as The Stull Act of California. This act requires yearly evaluation of educators and outlines the steps for dismissal of a certified educator. By 1973, twenty-seven other states had adopted similar laws (Borich & Fenton, 1977).

Studies conducted by the National Education Association in 1964 and 1969 (cited in Travers, 1973) revealed that an increasing number of school systems were using formal teacher evaluation procedures. However, improvements in the methods used were minimal as judged by the responding teachers. In 1969, over 90% of the teachers surveyed supported the idea of regular teacher evaluation.

The emphasis on accountability in education by communities and by educators encouraged educational changes. The "doing" aspect of the dual "knowing" and "doing" roles of teachers received emphasis in teacher education through the performance or competency-based education strategy. The Model Elementary Teacher Education

TM008 608

Programs, sponsored by the U.S. Office of Education in 1967, gave impetus to the evaluation of teaching competencies. Thus, assessment of teaching behaviors spurred new interest in evaluation (Dickson & Saxe, 1973).

Rosner (1972) contended that performance assessment should be just as stringent as knowledge assessment, and that teaching performance can be systematically evaluated. "The power of the idea resides in the clear expression of the objectives of the teacher education program and the ability to assess teacher competencies," according to Rosner (p. 34). He maintained that competency-based certification for preservice and inservice teachers should begin as soon as instruments for assessing their performance are developed. He said, "The preparation of instruments to define performance criteria is the sine quo non of competency-based certification" (p. 30).

General problems which confound the effectiveness of such instruments for the assessment of teacher competence have been identified by McNeil and Popham (cited in Travers, 1973). These problems involve the person selected to do the evaluation, the kind of instrument used, the appropriate use of the instruments, and proper application of data collected. Gallagher, Nuthall, & Rosenshine (1970) maintained that adequate training programs in the use of observation instruments can resolve some of these problems.

Although training in using assessment instruments seems essential, little research has been done in aspects of training

observers or assessors. Hylton and Quellmalz (1970) in inspecting Simon and Boyer's Mirrors for Behavior, which lists observation instruments, found only two descriptions of observer training programs. They concluded that "in spite of the increasing use of observation scales...one aspect of scale development remains singularly undocumented; that of observer training" (p. 53). A search of related literature revealed little or no reference to changes in this situation since 1970. Lack of such studies has implications for any organization involved in developing systems for observation of teachers and in training for use of the systems.

The State Board of Education of Georgia is currently involved in developing a system for assessing teacher competence as part of the state's teacher certification procedure. The University of Georgia received a grant from the State Department of Education in 1976 to develop lists of competencies and to develop instruments to be used in the assessment process. Dr. Charles E. Johnson is the director of this Student/Beginning Teacher Assessment Project (SBTAP). Five observation instruments have been developed (Okey et al., 1977) by the project staff. The project was also to develop training programs in the use of the instruments.

It was for these reasons, a specific need in the State of Georgia and lack of documentation in methods of assessor training, that this research was needed. The prupose of the study, therefore, was to compare the effectiveness of different methods of training assessors in the use of instruments for assessing teaching behavior.

The two instruments used in the study were developed by the SBTAP--the Classroom Procedures Assessment Instrument (CPI) and the Interpersonal Skills Assessment Instrument (ISI).

The methods of training were selected through practical and theoretical considerations. The large number of trainees used in the study and the prospective numbers to be trained later by the State, plus limited training time, made the use of on-the-spot, live classroom observation impossible. The training workshops were also held during August when school was not in session. A review of the literature related to training programs in education and other fields indicated that vidiotapes of classroom situations could be used in place of live observation. Research concerning protocol materials (Borg, Kelley, Langer, & Gall, 1970; Smith, Orlosky, & Borg, 1973) was supportive of the effectiveness of videotaped models for demonstrating teaching competencies and providing feedback for improving teacher performance.

Specific research in using vidiotape training in observation techniques has been reported. Janet Bower (1970) at Jersey City State College used videotaped sequences of classrooms as "the basis for instruction in techniques of observing and recording behavior" (p. 43). She found the video sequences to be useful for clarifying concepts and training preservice teachers to distinguish different forms of classroom behavior. Research by Morrison and Childs (1969) found that videocassettes could be used to improve

uniform understanding of concepts on rating forms by showing teaching performances. Hylton and Quellmalz (1970) used videocassette sequences for observer training to refine an observation system and to become familiar with the rating scale used. They found videotaped sequences to be of great value.

Programmed materials have widely been used in training programs. Rummler (1970) stated that "in areas where people have human performance problems to solve, and expect results for dollars invested, programmed instruction is flourishing" (p. 31). Crowder (1964) defined programmed instruction as "the organization of instructional materials in sequences optimally organized to produce student learning" (p. 28). Linear and branching programs are the best known forms of programmed material. Linear programs provide a sentence-by-sentence progression through content material by using a fixed sequence of small bits of information called frames. It is often called "Skinnerian Programming" for B.F. Skinner whose work with teaching machines and programming is widely known. Intrisic programming is a second form. It is often called a "branching program" as each learner response determines the next sequence of events. Thus, a learner may skip certain frames depending on his response (Markle, 1969).

Adjunct programming is a less expensive, often more relevant, and easier to write program than either branching or linear forms. It is essentially a series of questions which direct the learner to key points in a text or film. Such questions force the learner

to analyze the material. Research at Lackland Air Force Base indicated that adjunct programming can be just as effective as corresponding linear programs (Zaccaria & Adams, 1964). Interactive programming is a way of allowing learner involvement with materials presented for self-instruction. It usually combines two forms of media. Early films and filmstrips incorporated a series of questions for the learner to answer; this is still used as a form of adjunct, interactive programming. The latest application of interactive programming involves videocassettes which are a newer and more flexible medium for interaction (Barwick, 1973).

Programmed instruction gave new impetus to studies involving alternative student response modes. In reviewing research concerning student responses, Lumsdaine (1961) emphasized the importance of various student response modes on learning outcomes. Overt responses include those which are observable, such as answering orally, writing, marking an answer sheet, or pushing a lever on a teaching machine. Covert responses usually involve "thinking" or saying a response internally (Lumsdaine, 1961).

In reviewing research studies on the overtness of responses, Krumboltz (1961) concluded that the value of overt responding still remains questionable. Studies reviewed led him to hypothesize that overt responses may be more effective with difficult materials, but covert responses are more efficient with easy material. In a research report in 1962, Krumboltz and Weisman reviewed a study

they did involving 54 undergarduates who were assigned randomly to four groups to study test interpretation. One group wrote responses, one group 'mentally composed' responses, one group read the series of questions complete with answers, and the fourth section, a control group, worked with a different program. A 50-item test was administered at the end of the study period and another form of the test was given two weeks later. The three groups yielded no significant differences on the immediate test, but the written response group scored significantly better on the delayed test. The researchers concluded that overt responding appeared to increase delayed retention. Research reported by Silberman (1962) resulted in a different conclusion. He compared a writing response, a writing-and-vocalizing response, and a covert response. Those using the covert response program did better than the other groups. However, this and the other studies Silberman reviewed involved very short programs.

Much research has been reported on the role of questions in eliciting both overt and covert responses from learrers. Rothkopf's mathemagenic behaviors (1970), Anderson's attention process (1970), the advance organizers studied by Ausubel (1968), and Wittrock's generative learning model (1977) support the effectiveness of questions in organizing material for student learning. However, their research suggested the need for continuing study in the organization of questions, especially with non-print media.

The materials developed for the SBTAP training workshops were designed to evaluate three different interactive techniques

using videocassette materials and adjunct questions about the two instruments used, CPI and ISI. The ten videocassette sequences used were adapted from a series of protocol training films produced at the Universities of Colorado and Utah, the Far West Lab, and two films used with the Purdue Observation System. Each sequence showed a teacher working with a group of students. A variety of subject areas and grade levels were shown. The sequences were selected and matched to the indicators on the two instruments. The printed materials consisted of a series of questions and answers which presented the meaning of specific indicators on the instruments. These questions helped the trainee to rate the teaching situations presented via the videocassette sequences. These materials were available to each treatment group in the study in following ways:

Group 1: Oral Discussion

This group viewed the ten videocassette sketches and rated the teaching situation using the specified indicators from the CPI or ISI. An instructor gave the trainees feedback about their responses and then conducted a discussion of the instruments using a prepared list of questions and answers available only to the instructor.

Group 2: Directed Discussion

This group viewed the videocassette sequences and wrote answers to the list of questions about the teaching situations and the instruments; these same questions were used by each group. The instructor, who had a copy of the suggested answers, discussed the questions and answers with the trainees.

Group 3: Self-tutorial

This group saw the videocassette sketches, rated each sketch, and checked their own ratings from a manual provided by the instructor. They proceeded to answer the printed questions and check their answers by comparing them to suggested answers given in the manual. There was no discussion among the trainees, nor with the instructor.

These three treatment groups, plus a Control Group, served as the independent treatment variable in the study. Grade level of certification of the trainees, either elementary grades K-5 or secondary grades 6-12, was the second independent variable.

The dependent variables studied were the answers given on an attitude questionnaire about the method of training and the scores the trainees received for rating Classroom Teacher Performance Lessons using the two instruments. These Classroom Teacher Performance Lessons were a 30-45 minute lesson on a videocassette taught by a student teacher or a beginning teacher in Clarke County. Each cassette was filmed by the University of Georgia's Instructional Resource Center under the direction of the staff of the SBTAP. Each lesson had been viewed by a panel selected by the project staff and rated using the CPI and ISI. These ratings formed a standard profile for each lesson to be used in giving a score to the trainees on their ability to rate the Teacher Performance Lesson using the two instruments.

The scores depended on the trainee's ability to rate the lesson as compared to the ratings on the standard profile. The profile

consisted of a center line on a graph representing the mean rating given by the panel for the lesson on each competency on the instruments. Six additional pairs of lines representing variation in the panel's judgment completed the graph. Each pair represented an increment of .5 standard error of the measure for a competency. If a trainee's calculated competency mean fell beyond the widest lines, representing three standard errors of the measure, the rating was unacceptable. If it fell within the widest lines, the trainee earned one point toward the total scores. If it fell within the second pair of lines, two points were earned; the third pair, three points; fourth pair, four points; fifth pair, five points. If the mean fell within the pair of lines nearest the composite mean, the score was six points for that competency rating. The two scores received consisted of the sum of points earned for each of the eight competencies on the CPI and the sum of points for the four competencies on the ISI, possible scores of 48 and 24, respectively.

A semantic differential questionnaire was used to assess the attitudes of the trainees toward their method of instruction. The questionnaire asked questions about the entire workshop with specific questions related to the instructional materials and processes used in training in the use of the CPI and ISI. The questionnaire was scored using a 5-point scale with five being the most positive attitude and one being the most negative.

The following were the hypotheses tested:

(1) There is no difference among the mean scores of the trainees by treatment group on the CPI ratings.

(2) There is no difference among the mean scores of the trainees by treatment group on the ISI ratings.

(3) There is no difference among the mean scores of the trainees by grade level of certification for the CPI.

(4) There is no difference among the mean scores of the trainees by grade level of certification for the ISI.

(5) There is no difference among the trained groups in the trainees' attitudes toward their instruction.

(6) Effects of treatment on mean rating ability scores and observer attitudes are consistent across the grade level of certification of the trainees.

(7) There is no difference among the mean scores of the trainees by treatment group (or by grade level of trainees) on the CPI rating of a lesson selected by grade level after additional practice.

(8) There is no difference among the mean scores of the trainees by treatment group (or by grade level of trainees) on the ISI rating of a lesson selected by grade level after additional practice.

The experimental population consisted of 69 trainees selected by the Clarke County School System to attend training workshops on the use of the five assessment instruments developed by the SBTAP. The trainees were assigned by the school system to the workshops,

51 to the first workshop held the first week in August and eighteen to the second workshop held two weeks later. Both workshops were conducted at Alps Road School in Athens, Georgia.

The 51 trainees in the first workshop were randomly assigned using a table of random numbers to each of the three treatment groups described earlier, seventeen in each group. The eighteen attending the second workshop served as a Control Group. The elementary and secondary teachers in each group were almost equally divided, 7-10, 9-8, 9-8, 8-10, respectively for the four groups. By years of teaching experience the groups were fairly evenly matched with the majority in each group falling within the 7-18 years of experience groups. The majority of trainees in each group had a masters degree or above. By gender, women outnumbered the men in each group, 13-4, 14-3, 14-3, 10-8, respectively.

The training workshop was a week long. Eight hours of training were used to instruct the trainees in using the CPI and ISI on the first two days of the workshop. The first Classroom Teacher Performance Lesson was then rated. The next two days of the workshop focused on the other assessment instruments developed by the SBTAP with time for two practice sessions using the CPI and ISI. On Friday a second Classroom Teacher Performance Lesson was rated; this criterion lesson was selected by grade level, either K-5, 6-8, or 9-12, to match the level taught by the trainee. The attitude questionnaire was also completed on Friday. The Control Group, enrolled in a later workshop, rated the first Classroom Performance Lesson prior to any training.

Three instructors worked in the training sessions. Each instructor had previously worked with the training materials. One instructor monitored the Self-tutorial Group. The other two instructors worked with the trainees in Groups 1 and 2, spending one-half of the training sessions with each group and then rotating.

All hypothesis testing was done using analysis of variance procedures. Where significant F-ratios were found Newman-Keuls procedures were done to locate the source of differences. The analysis of variance procedures were done using a computer program from the Statistical Package for the Social Sciences (1975). The ten percent level of probability was judged to be an appropriate level of statistical significance for the rejection of the null hypotheses.

Results and Conclusions

On the first criterion rating of a classroom Performance Lesson, the analysis of variance for the CPI rating indicated no significant difference between the elementary and secondary groups. The hypothesis of no difference by grade level was not rejected. There was a significant difference by treatment group, $F(3,61) = 3.858$, $p < .014$. The hypothesis of no difference by treatment group was rejected. The Newman-Keuls Procedure identified significant differences between Group 2, the Directed Discussion Group, and the Control Group and between the Self-tutorial Group and the Control Group (See Tables 1 and 2). The group means for groups 1-4 were 37, 40.88, 39.77 and 35.17, respectively out of a possible 48.

These findings indicated that training using written responses with either oral discussion or with printed feedback is more successful than just oral discussion or no treatment for the CPI. The oral discussion mean was higher than the control group, but not significantly so.

For the first criterion rating on the ISI the analysis of variance indicated no significant difference by grade level of certification. The hypothesis of no difference by grade level was not rejected. A significant difference by treatment group was found $F(3,61) = 2.892$, $p < .042$, thus rejecting the hypothesis of no difference by treatment group on the ISI rating. The Newman-Keuls procedure revealed that the significant differences were between groups 1 and 4 and between groups 2 and 4 (See Tables 3 and 4). The means were 15.71, 15.53, 14.47, and 12.34, in order by group.

The more subjective percentage rating on the ISI might account for the higher ratings for Group 1 and 2 both of which involved oral discussion of the instrument during training.

For the combined scores, Group 2 had the highest mean (56.4) followed by Groups 3 (54.24), 1 (52.71), and 4 (47.50). Each group did significantly better than the Control Group, $F(3,61) = 4.418$, $p < .007$. This indicated that training did make a difference. Since the self-tutorial group scored the second highest and was not significantly different from groups 1 and 2, it would seem to be the best procedure for a workshop having limited instruction time or money to hire instructors. It could be followed independently freeing the workshop instructor to work with small groups in other aspects of training (See Tables 5 and 6).

For the second or delayed rating on the CPI and ISI by the 51 trainees there were no significant differences among treatment groups on the CPI or ISI or combined scores. On the CPI rating the elementary trainees scored significantly higher than the secondary trainees $F(2, 45) = 5.919$, $p < .018$. For the combined scores the elementary trainees also scored higher $F(2, 45) = 4.390$, $p < .04$. (See Tables 7 and 8).

Additional practice and rating in a trainees own grade level area appeared to diminish differences among the treatment groups. The reversal of findings by grade level was interesting. Evidently, elementary trainees are better at assessing elementary lessons, than secondary trainees are at assessing high school teachers.

Responses on the attitude questionnaire for the questions about the CPI and ISI training yielded no significant differences $F(1, 50) = .702$, $p < .99$. Thus, the hypothesis was rejected.

Implications

Treatment did make a difference but the effects of a specific treatment dissipated with additional practice and rating within the trainee's own grade level. However, the result by grade level is an area needing further research to determine if trainees should only rate within their own grade level. The findings regarding the self-tutorial approach offer practical implications for certain training situations.

Another area which should be investigated is to see if differences are found when the trainees assess real or live situations as compared to videocassette lessons.

TABLE 1

Analysis of Variance for the CPI Rating of First Classroom Performance Lesson

Source of Variation	df	MS	F	P
Treatment Group (V1)	3	119.79	3.86	.014
Certification (V2)	1	6.06	0.20	.999
VI x V2	3	21.09	0.68	.999
Within	61	31.05		
TOTAL	68	34.08		

$\alpha = .10$

TABLE 2

Newman-Keuls: CPI, First Rating

$\bar{X}_1 = 37.00$	$\bar{X}_2 = 40.88$	$\bar{X}_3 = 39.77$	$\bar{X}_4 = 35.17$
$\bar{X}_2 - \bar{X}_4 = 5.71$	$q_{4,60} = 4.23$		$p < .10$
$\bar{X}_3 - \bar{X}_4 = 4.60$	$q_{3,60} = 3.40$		$p < .10$
$\bar{X}_2 - \bar{X}_1 = 2.88$	$q_{3,60} = 2.87$		$p > .10$
$\bar{X}_1 - \bar{X}_4 = 1.83$	$q_{2,60} = 1.61$		$p > .10$

NOTE: X_1 is mean for the Oral Discussion Group; X_2, Directed Discussion Groupl X_3, Self-tutorial Group; X_4, Control Group. This is the same notation for each Newman-Keuls table.
$\alpha = .10$.

TABLE 3

Analysis of Variance for First ISI Rating

Source of Variation	df	MS	F	P
Treatment Group V1)	3	43.06	2.89	.042
Certification (V2)	1	20.33	1.37	.246
V1 x V2	3	25.85	1.74	.167
Within	61	14.89		
TOTAL	68	16.67		

$\alpha = .10$.

TABLE 4

Newman-Keuls: ISI, First Rating

$\bar{X}_1 = 15.71$	$\bar{X}_2 = 15.53$	$\bar{X}_3 = 14.47$	$\bar{X}_4 = 12.34$
$\bar{X}_1 - \bar{X}_4 = 3.37$	$q_{4,68} - 3.59$		$p < .10$
$\bar{X}_2 - \bar{X}_4 = 3.19$	$q_{3,68} = 3.41$		$p < .10$
$\bar{X}_3 - \bar{X}_4 = 2.13$	$q_{2,68} = 2.28$		$p > .10$
$\bar{X}_1 - \bar{X}_3 = 1.24$	$q_{3,68} = 1.33$		$p > .10$

$\alpha = .10$

TABLE 5

Analysis of Variance for Combined CPI/ISI, First Rating

Source of Variation	df	MS	F	P
Treatment Group (V1)	3	261.19	4.42	.007
Certification (V2)	1	48.58	0.82	.999
V1 x V2	3	71.39	1.21	.315
Within	61	59.12		
TOTAL	68	68.09		

TABLE 6

Newman-Keuls: Combined CPI/ISI

$\bar{X}_1 = 52.71$	$\bar{X}_2 = 56.41$	$\bar{X}_3 = 54.24$	$\bar{X}_4 = 47.50$
$\bar{X}_2 - \bar{X}_4 = 8.91$	$q_{4,68} = 4.78$	$p < .10$	
$\bar{X}_3 - \bar{X}_4 = 6.74$	$q_{3,68} = 3.61$	$p < .10$	
$\bar{X}_1 - \bar{X}_4 = 5.21$	$q_{2,68} - 2.79$	$p < .10$	
$\bar{X}_2 - \bar{X}_1 = 3.70$	$q_{3,68} = 1.99$	$p > .10$	

TABLE 7

Analysis of Variance: Second Rating

Source of Variation	df	MS	F	P
	Combined CPI/ISI			
Treatment Group (V1)	2	2.95	.07	.999
Certification (V2)	1	200.00	4.39	.040
V1 x V2	2	4.10	.09	.999
Within	45	45.55		
TOTAL	50	45.45		
	CPI Total			
V1	2	5.81	.276	.999
V2	1	124.50	5.920	.018
V1 x V2	2	6.13	.291	.999
Within	45	21.03		
TOTAL	50	22.09		
	ISI TOTAL			
V1	2	.92	.06	.999
V2	1	8.91	.60	.999
V1 x V2	2	.50	.03	.999
Within	45	14.96		
TOTAL	50	13.68		

TABLE 8

Means of Trainees on the Second Rating

Certification Level	CPI/ISI Total	CPI	ISI
Elementary	59.16	40.36	18.80
Secondary	55.12	37.12	18.00

REFERENCES

Anderson, R.C. Control of student mediating processes during verbal learning and instruction. Review of Educational Research, 1970, 40, 349-70.

Ausubel, D.P. Educational psychology: A cognitive view. New York: Holt, Rinehart and Winston, 1968.

Barwick, J., Kranz, S., & Beker, G. The compleat videocassette users guide: Principles and programming. White Plains, N.Y.: Knowledge Industry Publication, Inc., 1973.

Borg, W.R., Kelley, M., Langer, P., & Gall, M. The minicourse: A microteaching approach to teacher education. Ceverly Hills: MacMillan, 1970.

Borich, C., & Fenton, K. The appraisal of teaching: Concepts and process. Reading, Mass.: Addison-Wesley, 1977.

Bowe[illegible], J. Using videotape in a foundations course in teacher education. Audiovisual Instruction, 1970, 15, 43-44.

Crowder, W. Programmed instruction compared with automate instruction. In G. Ofiesh & W. Meierhenry (Eds.), Trends in programmed instruction. Washington: National Education Association, 1964.

Dickson, G.E., Saxe, R.W., et al. Partners of educational reform and renewal. Berkeley: McCutchan Publishing Corporation, 1973.

Gallagher, J., Nuthall, G., & Rosenshine, B. Classroom observation. Chicago: Rand McNally, 1970.

Hylton, J.A. & Quellmalz, E. Development of an observer training videotape sequence. Educational Technology, 1970, 10, 53-55.

Krumboltz, J. Meaningful learning and retention: Practice and reinforcement variables. Review of Educational Research, 1961, 31 535-46.

Krumboltz, J., & Weisman, R.G. The effect of overt versus covert responding to programmed instruction on immediate and delayed retention. Journal of Educational Psychology, 1962, 53, 89-92.

Lumsdaine, A.A., & Glaser, R. (Eds.) Teaching machines and programmed learning: A source book. Washington: National Education Association, 1960.

Markle, S. Good frames and bad: A grammar of frame writing (2nd ed.), New York: John Wiley and Sons, 1969.

Morrison, V.B., & Childs, J. Strategies for the application of videotape tape in teacher education. Audiovisual Instruction, 1969, 14, 43-48.

Nie, N., et al. Statistical Package for the Social Sciences (2nd ed.). New York: McGraw-Hill, 1975.

Okey, J.R., Long, J., Capie, W., & Johnson, C. Teacher performance assessment instruments: Manual for data collectors. Athens, Georgia: University of Georgia, 1977.

Rothkopf, E.Z. The concept of mathemagenic behaviors. Review of Educational Research, 1970, 40, 325-336.

Rosner, B. The power of competency-based teacher education. Boston: Allyn & Bacon, 1972.

Rummler, G.A. PI-where the action is. Educational Technology, 1970, 10, 31.

Silberman, H. Self teaching devices and programmed materials. Review of Educational Research, 1962, 32, 179-193.

Smith, B.O., Orlosky, D., & Borg, G. Handbook on the development and use of protocol materials for teacher education. Chipley, Fla.: Panhandle Area Education Cooperative, 1973.

Travers, R.M.W. (Ed.) Second handbook of research on teaching. Chicago: Rand McNally, 1973.

Wittrock, M.C. A model of human generative learning. (Paper presented at the 1977 American Educational Research Association's Annual Meeting, New York).

Zaccaria, M.A., & Adams, C.F. Adjunct programming. In G. Ofiesh & W. Meierhenry (Eds.), Trends in programmed instruction. Washington: National Education Association, 1964.